

移动场景扩展现实超高清技术白皮书

Technical White Paper on UHD Extended Reality in Mobile Scenarios

UHD World Association
世界超高清视频产业联盟



前 言

本文件由UWA联盟虚拟现实工作组组织制订，并负责解释。

本文件发布日期：2024年12月31日。

本文件由世界超高清视频产业联盟提出并归口。

本文件归属世界超高清视频产业联盟。任何单位与个人未经联盟书面允许，不得以任何形式转售、复制、修改、抄袭、传播全部或部分内容。

本文件主要起草单位：

中央广播电视总台、华为技术有限公司、咪咕文化科技有限公司、深圳创维数字技术有限公司、北京字节跳动网络技术有限公司、中国电子技术标准化研究院、中国信息通信研究院、杭州微帧信息科技有限公司

本文件主要起草人：

李岩、张钰、郎凤岐、周子丰、赵海龙、朱家悦、杨友庆、毕蕾、陈文斌、陈春明、张月川、胡超、耿一丹、徐遥令、翟云、陈曦、汪立民

免责声明：

- 1, 本文件免费使用，仅供参考，不对使用本文件的产品负责。
- 2, 本文件刷新后上传联盟官网，不另行通知。

目 录

1. 移动场景扩展现实超高清技术概述	1
2. 移动场景扩展现实超高清关键技术	1
2.1 内容生产	2
2.2 内容编码	6
2.3 内容传输	8
2.4 内容播放	9
3. 移动场景扩展现实的应用场景	11
3.1 消费者场景	11
3.1 企业场景	15
4. 移动场景扩展现实面临的挑战	20
4.1 沉浸式体验	20
4.2 丰富的内容	22
4.3 便携和续航	22
5. 技术标准化与产业政策建议	23
5.1 关键技术标准化及发展方向	23
5.2 产业政策建议	25
6. 发展趋势展望	26
7. 附录	29
7.1 缩略语	29
参考文献	30

1. 移动场景扩展现实超高清技术概述

移动场景扩展现实超高清技术是指在移动蜂窝网络下使用扩展现实设备（XR）的超高清技术，XR涵盖了虚拟现实（VR）、增强现实（AR）和混合现实（MR）等多种形式、应用场景包含个人消费者场景及企业用户场景。

5G自2019年正式商用以来，截至2023年全球已拥有超过15亿用户，提供了远高于4G的用户体验，5G技术同时也在推动制造业、能源、公共事业、医疗保健、媒体和娱乐等行业的转型，成为迄今为止增长最快的移动技术。截至2024年上半年中国累计建设近375万个5G基站，基本完成全国覆盖。2024年5G-A的第一个标准版本3GPP Rel-18正式冻结，标志着5G-A进入商用元年，全球领先的13大运营商联合宣布成为5G-A网络先锋，中国移动更是率先在杭州启动了世界上第一个商业5G-A网络，并计划在年底前扩展到全国300多个城市，标志着5G-A从技术验证步入商用部署阶段。

国际市场在2024年上市了其首款混合现实（MR）头显设备，该设备主打超高清观影体验。采用Micro OLED屏幕，单眼分辨率达到4K以上，并且支持8K超高清解析度，为用户提供了出色的图像质量，它还集成了流媒体服务，支持高质量的8K内容播放，为用户带来沉浸式的观影体验，让观看超高清电影真正走向了移动场景。与此同时，整个产业链都在积极跟进，高通发布了支持单目4K，双目8K的芯片，国内创维、玩出梦想等终端厂家也发布了双目8K MR样机，咪咕上线咪咕视频-Vision等软硬件产品。

5G/5G-A快速的发展为移动场景下的超高清视频提供了强大的网络支持，扩展现实超高清视频技术则在显示和音频方面带来了革命性的体验，共同推动了移动场景下超高清视频的发展，为超高清技术发展提供了新的发展方向。

2. 移动场景扩展现实超高清关键技术

随着5G/5G-A持续演进和创新，将提供更高的带宽、更低的时延，移动网络下扩展现实设备观看超高清视频将进入快车道，对超高清内容生产、编码、传输、播放等整个端到端的技术都带来新的挑战。超高清视频包含MPEG、MV-HEVC等多种格式，包含双目3D视频、全景视频、体积视频等多种形态，本白皮书将基于近期发展快且已经有商用高码率8K视频发布的MV-HEVC格式双目3D视频进行介绍。



图1 移动场景扩展现实超高清内容系统

2.1 内容生产

移动网络场景下的超高清技术相对传统平面超高清技术为用户提供了更加丰富和沉浸的观看体验，这不仅需要更高的分辨率和帧率，同时还要考虑全景拍摄、3D效果、用户交互和多视角切换等问题，通常需要专门的软硬件来支持，随着空扩展现实类产品的发布，空间视频和空间音频开始受到广泛关注。

1、空间视频的定义：左、右眼分别传送不同的视图，会带来更丰富的用户体验，向用户的左眼呈现一幅图像，同时右眼呈现另一幅相关的图像，以产生立体视觉效果，大脑同时接收来自两只眼睛的视觉刺激而产生的深度感知，空间视频能够存储除视频本身的额外深度信息。

2、空间视频关键技术及产品演进：为提升空间视频体验，目前在技术上可以提供XR平台类工具套件，为开发者提供激光测量（LiDAR）场景重建、人体运动捕捉跟踪、垂直和水平平面检测、图像检测、3D对象检测、3D对象扫描、4K HDR视频捕捉等底层技术，利用MV-HEVC（Multiview High Efficiency Video Coding）的视频编码可显著提升3D视频观看体验。用户还可以对某一真实物体进行扫描建模，并通过对象追踪功能（Object Tracking）实时精准定位物体位置。部分产品在功能上可以通过对象捕捉（Object Capture）进行3D模型创建，实现空间视频录制。使用更新的空间音频功能，可根据个人的身体生理结构定制出不同的空间音频效果。

3、非专业的空间视频制作流程：随着科技的发展，日常空间视频制作已经开始逐渐普及并且操作更加简单。一些产品可以基于人类的双目立体视觉原理，通过手机主摄像头与超广角摄像头模仿人眼采集具有一定水平视差的双目画面，生成MV-HEVC编码视频。空间视频拍摄方式和常规二维视频拍摄方式类似，用户打开手机相机APP，选择视频模式并横置，点击空间视频图标符号，点击录制按钮即可拍摄空间视频。

行实时的协调合作。

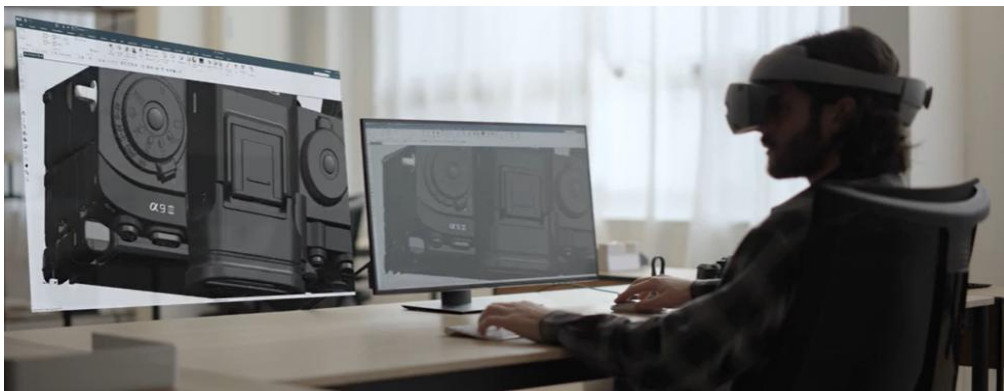


图5 NX Immersive Designer

5、空间视频呈现及交互：目前各厂商已陆续推出各类沉浸式视频短片。相比传统电影的视野范围（40°-60°），这种沉浸式视频可以提供180°视角的8K画面，同时配合耳机的空间音频功能，整个观看体验更加身临其境。结合佩戴设备上8K分辨率显示器和高效眼追踪系统，将MV-HEVC编码视频分别呈现给人的左眼和右眼，最终产生具有空间立体感的3D画面，同时通过交互功能，用户与空间视频画面进行实时互动。当前咪咕视频、腾讯视频等多个内容制造商开始出品此类节目，多赛多屏同看（咪咕视频-Vision）、王者荣耀沙盘（腾讯视频-Vision）等交互场景已开始应用。



图6 多赛多屏同看功能

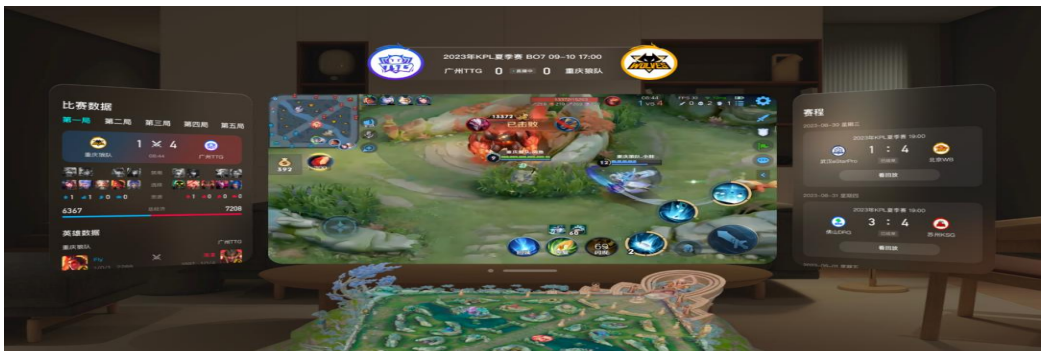


图7 王者荣耀沙盘功能

6、空间视频生产设备发展情况：目前，iPhone15Pro及以上型号、Vision Pro、Quest Pro、Quest3、PICO4 Ultra等新型终端设备大幅提升硬件运算能力，开始支持空间视频和空间音频，并支持拍摄和预览MV-HEVC编码标准的空间视频，让空间视频制作逐步实现便携化。

表1 支持空间视频制作的主流头显设备参数对比

	参数	Apple Vision Pro	PICO 4 Ultra	Meta Quest 3	Quest Pro
近眼显示	屏幕	2x1.42"Micro-OLED(内屏) OLED(外屏)	2x2.56"Fast-LCD	2x2.56"Fast-LCD	2x2.48"Fast-LCD
	单眼分辨率	3680x3140	2160x2160	2064x2208	1800×1920
	光学方案	3P Pancake	Pancake	Pancake	Pancake
	FOV	100°	105°	110°	106°
	PPD	34	20.6	19.6	16.9
	刷新率	90/96/100Hz	72/90Hz	72/80/90/120Hz	72/90Hz
	HDR	支持	不支持	不支持	不支持
	瞳距调节	51-75mm	58-72mm	53-75mm	55mm-75mm
空间交互	操作系统	Vision OS	Android OS	Horizon OS	Horizon OS
	追踪方案	6DoF 空间定位方案，Inside-Out(12颗摄像头)	6DoF 空间定位方案，Inside-Out	6DoF 空间定位方案，Inside-Out(4颗摄像头)	6DoF 空间定位方案，Inside-Out
	手柄	无	无光环 6DoF 手柄	无光环 6DoF 手柄	无光环 6DoF 手柄
	手势	支持	支持	支持	支持
	摄像头数量	12 (2300 万像素 RGB 主相机×2，下视角相机×4，外侧视角相机×2，眼球追踪红外相机×4)	7 (3200 万像素 RGB 摄像头×2；iToF 深度感知摄像头×1；环境追踪摄像头×4)	4 (黑白摄像头×2；400 万像素 RGB 摄像头×2)	10 (红外摄像头×2；1600 万像素 RGB 摄像头×1)
	传感器数量	5 (原深感测镜头；LiDAR 感测器；四个惯性测量单元(IMU)；闪烁感测器；环境光线感应器)	13 (1 颗 IMU 传感器；12 颗红外传感器)	4 (重力感应器；距离传感器；加速度感应器；陀螺仪感应器)	4 (重力感应器；距离传感器；加速度感应器；陀螺仪感应器)
	面部追踪	支持	不支持	不支持	支持
	眼动追踪	支持	不支持	不支持	支持
	手部追踪	支持	支持	支持	支持

	空间视频制作	支持	支持	支持	支持
性能续航	芯片	Apple M2 (8 核 CPU、10 核 GPU、16 核神经网络引擎) Apple R1(12ms 延迟, 256GB/s 内存带宽)	高通骁龙 XR2 Gen2 (两个 2.05GHz 核心、四个 2.36GHz 核心)	高通骁龙 XR2 Gen2 (两个 2.05GHz 核心、四个 2.36GHz 核心)	高通骁龙 XR2 Gen1
	RAM	16GB	12GB	8GB	12GB
	ROM	256GB/512GB/1T	256GB	128GB/512GB	256GB
	电池容量	外接 3166mAh	内置 5700mAh	内置 4879mAh	内置 5000mAh
其他	头显重量	600-650g (电池重 353g)	含顶部绑带 585g	含肩带和面部接口 515 克	722g
	WIFI	Wi- Fi6	Wi-Fi7	Wi-Fi6E	Wi- Fi6
	扬声器	支持空间音频	双立体扬声器, 支持 3D 空间音效	双立体扬声器	双立体扬声器
	麦克风	6 个	4 个, 支持空间音频录制	2 个	3 个
	发布时间	2023 年 6 月 6 日	2024 年 8 月 20 日	2023 年 10 月 10 日	2022 年 10 月 12 日

2.2 内容编码

当前移动场景扩展现实设备超高清视频主流编码方式为多层格式高效视频编码（MV-HEVC/3D-HEVC）。

1、MV-HEVC/3D-HEVC定义：不同于传统的AVC/HEVC单层3D上下/左右方式，MV-HEVC采用多层编码设计，通过利用同一帧左右眼视图之间的冗余信息做帧内预测，实现更高的压缩率。这种格式的优点是可以向下兼容，对于不支持MV-HEVC解码的平台，只解码base view（单眼），仍然能够以2D形式观看。3D-HEVC是一种更高级的3D视频扩展格式，由多个view和关联的depth maps编码表示，与MV-HEVC相比，通过指定新的块级视频编码工具，利用视频纹理和深度之间的相关性进一步提高压缩率。编码器多layer间关系如图8所示，实线为MV-HEVC层间参考关系（没有depth layer），虚线为3D-HEVC层间参考关系。

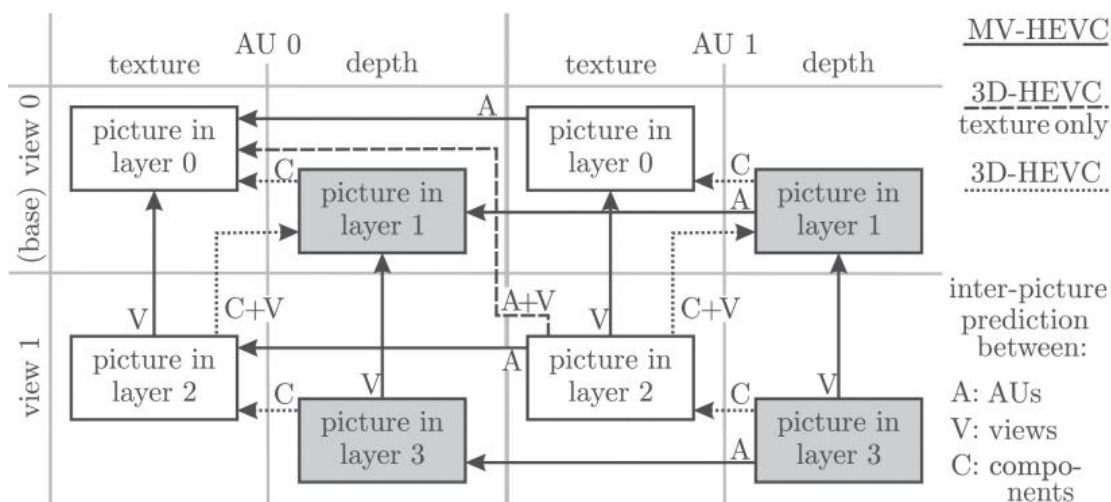


图8 MV-HEVC/3D-HEVC格式示意图

2、编码预测：

1) MV-HEVC Inter-Layer预测：MV-HEVC 架构的一个关键优点是它不会更改slice以下HEVC单层编码所需的语法或解码过程，可以重用现有实现，无需对HEVC解码器进行大的更改。与传统HEVC基于时域的帧间预测不同，MV-HEVC还允许从同帧的不同view进行预测，称为inter-view预测。相同view的时域参考帧和相同时时间的inter-view参考帧都可以出现在参考帧列表的任何位置。

2) 3D-HEVC Inter-Layer预测：3D-HEVC在MV-HEVC的基础上提供了更加丰富的inter-layer预测：a) 结合的时域和IV预测（A+V），即不同view之间的时域帧间参考；b) component间预测（C），相同AU和view中不同component之间参考；c) 结合的component间预测和IV预测（C+V），即相同AU中不同view和不同component间参考。

3、压缩：3D SBS/TB, MV-HEVC, 3D-HEVC三种3D格式压缩性能对比如表2，MV-HEVC对比普通3D上下/左右，压缩效率平均提升32%，3D-HEVC对比MV-HEVC，压缩率进一步提升15%。

表2 MV-HEVC/3D-HEVC压缩率对比

1) Two-view texture (stereo)		2) Three-view texture and depth		
Test set	MV vs. Sim. ¹	Test set	3D vs. Sim. ¹	3D vs. MV
		<i>PoznanHall</i>	-40.5% ⁴	-19.3% ⁴
		<i>PoznanStreet</i>	-41.5% ⁴	-11.9% ⁴
<i>UndoDancer</i>	-37.4% ²	<i>UndoDancer</i>	-50.9% ⁴	-19.2% ⁴
<i>GTFLy</i>	-42.6% ²	<i>GTFLy</i>	-55.3% ⁴	-18.5% ⁴
<i>Bmx</i>	-29.0% ²	<i>Balloons</i>	-40.3% ⁴	-19.3% ⁴
<i>Band06</i>	-34.7% ²	<i>Newspaper</i>	-41.0% ⁴	-19.9% ⁴
<i>Musicians</i>	-25.3% ²	<i>Kendo</i>	-40.0% ⁴	-21.4% ⁴
<i>Poker</i>	-24.9% ²	<i>Shark</i>	-59.2% ⁴	-24.4% ⁴
Average	-32.3% ²	Average	-46.1% ⁴	-19.3% ⁴
Enh. Only	-70.8% ³	Enh. Only	-72.8% ⁵	-34.3% ⁵

¹Simulcast coding with single-layer HEVC. Savings calculated based on:

²T₀, T₁ (PSNR and bit rate); ³T₁ (PSNR and bit rate);

⁴six synthesized views (PSNR) and T₀, D₀, T₁, D₁, T₂, D₂ (bit rate);

⁵six synthesized views (PSNR) and D₀, T₁, D₁, T₂, D₂ (bit rate).

2.3 内容传输

2017年至2023年期间，3GPP发布了三个版本的5G技术标准，Release-15版本中，主要确定了5G的物理层和数据链路层技术，包括频谱分配、调制编码和传输技术等。Release-16版本中，进一步完善了5G的控制层技术，包括网络切片、边缘计算和无线资源管理等。Release-17版本中，主要关注5G的增强技术，如URLLC和mMTC等，2024年6月，3GPP正式发布Release-18，标志5G-Advanced（5G-A）阶段启动。

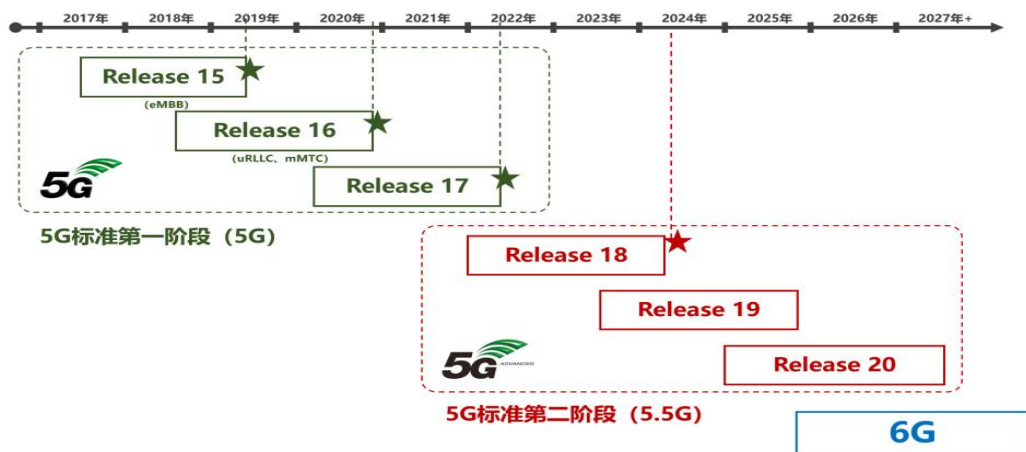


图9 3GPP的5G标准各版本发布时间表

表3 3GPP的5G/5G-A标准参考表

关键性能指标	5G	5G-A
--------	----	------

下行速率	1Gbps	10Gbps
上行速率	100Mbps	1Gbps
时延及可靠性	1 个/m2	10 个/m2
感知/定位精度	米级	感知：广域亚米级； 定位：厘米级

网络带宽是保障超高清视频流畅播放和即点即开的关键因素，经过华为实验室实测，当网络带宽达到视频码率的2倍时能够满足流畅播放需求，当网络带宽达到5倍码率时可满足即点即开体验需求，4K和8K的网络要求如表4所示：

表4 移动场景超高清视频保障带宽要求

分辨率	码率 (Mbps)	帧率 (Frame)	流畅体验带宽 (Mbps)	即点即开带宽 (Mbps)
4K	8~22.5	30/60	16~45	40~110
8K	20~85	30/60	40~170	100~425

2.4 内容播放

播放分为内容解析、内容解码、内容播放3个阶段，并且当前业界对视频播放已MOS分标准。

1、内容解析：1) 分离视角数据，HTM解码器会将原始数据分解为不同的视角层，例如layer0和layer1，这些层分别对应不同的视角；2) 不支持 MV-HEVC 解码的设备：只能识别 hvcC无法解析 lhvC 的语义，那么会创建普通 HEVC 解码器，仅解码主视角帧（main view），以实现单目兼容解码播放。

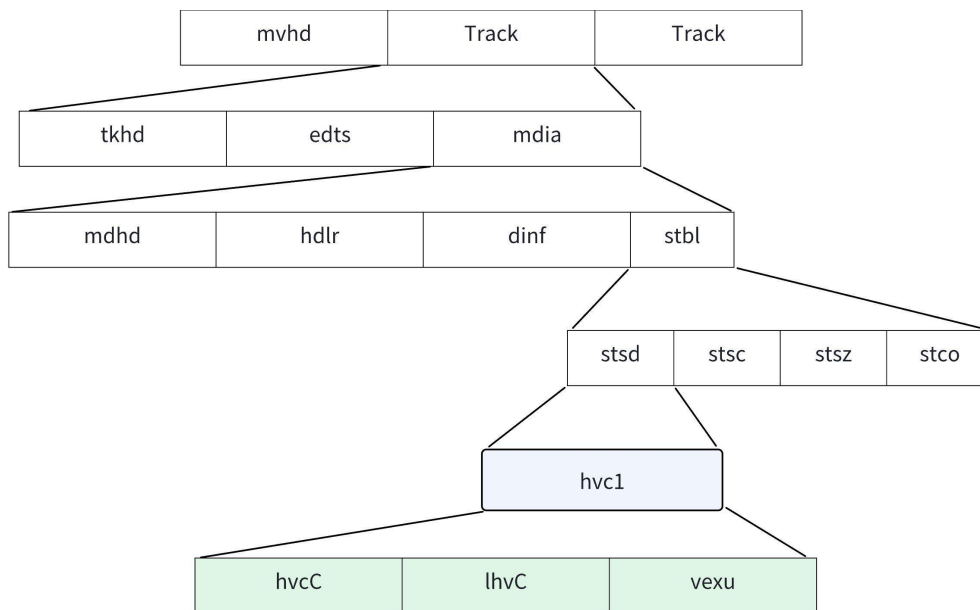


图10 MV-HEVC/3D-HEVC解析流程

- 2、内容解码：MV-HEVC的解码流程涉及多个步骤，以确保正确地处理和重建多视角视频内容：1) 获取解封装信息：首先，需要从封装文件中提取必要的信息，如lhvC信息。这一步是为了确保解码器能够正确解析MV-HEVC的二进制数据流。2) 使用HTM解码器进行解码：将提取的二进制数据送入HTM（High Efficiency Video Coding Multi-view）解码器进行处理。HTM解码器能够处理MV-HEVC的多层编码结构，生成多个视角的原始数据。
- 3、内容播放：与普通 2D 视频播放相比，MV-HEVC 3D 视频播放主要差异点：1) 视频解析器除了需解析音频流、视频流以外，还需要解析动态元数据流（视差参数）；2) 视频渲染器需利用视差参数构建渲染矩阵，将视频解码帧渲染到上屏帧缓冲中，具体播放流程见图11。

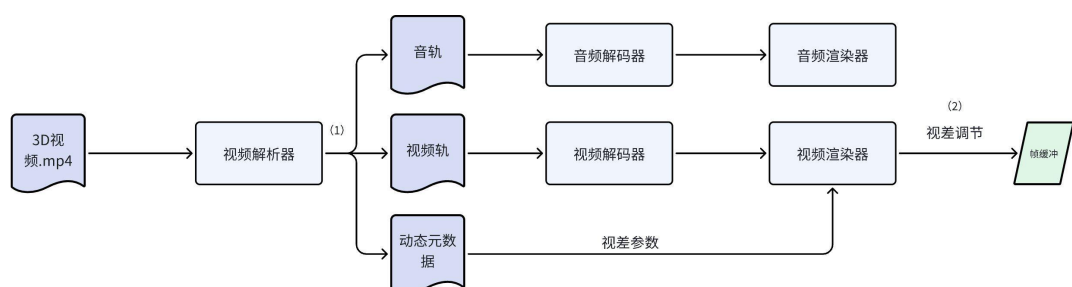


图11 MV-HEVC/3D-HEVC视频播放流程

- 4、内容播放的评价：与传统平面视频相似，MV-HEVC视频通过MOS 主观评分来进行，在清晰度、色彩、亮

度、噪声及伪影这四个维度的基础上，增加了“3D合像”维度，用于衡量视差调节是否能达成人眼舒适合像的目标，以 5 分制 MOS 分为例：

表5 移动场景扩展现实超高清视频播放MOS打分逻辑

1 分	2 分	3 分	4 分	5 分
呈现质量：合像困难、沉浸感非常差 交互质量：交互时延非常长	呈现质量：合像正常但能察觉到明显疲劳，沉浸感差且能感觉到明显出镜 交互质量：交互时延长且感觉到明显延迟	呈现质量：合像正常但能察觉到疲劳，沉浸感差且能感觉到出镜 交互质量：交互时延长且能感觉到延迟	呈现质量：合像正常且几乎察觉不到疲劳，沉浸感好几平像身临其境 交互质量：交互时延短且几乎感觉不到延迟	呈现质量：合像正常且舒适，沉浸感好且身临其境 交互质量：交互时延短且感觉不到延迟

3. 移动场景扩展现实的应用场景

3.1 消费者场景

移动场景扩展现实技术通过将虚拟信息和现实环境相结合，为个人消费者用户提供丰富的交互体验，当前在大空间场景和车载场景已有应用。

1. XR大空间场景

XR大空间支持多人在同一虚拟环境中进行体验，具有更强的交互性以及更广阔的交互环境，通常采用光学动作捕捉等高精度定位技术，以实现对用户空间中位置和动作的精确追踪，被广泛应用于文化和娱乐领域，如博物馆、主题公园等，通过VR技术重现历史场景或提供沉浸式的故事体验。

案例1：《琴岛奇遇》XR大空间

2024年8月，由厦门市鼓浪屿-万石山风景名胜区管理委员会监制、咪咕依托5G-A技术优势，整合“VR多人互动+沉浸体感体验”内容，打造了首个鼓浪屿主题VR大空间《琴岛奇遇》，通过“八闽文脉·记忆”“乐动

琴岛·守护” “钟声依旧·时光” “以乐会友·圆满” 4K超清电影级内容篇章，生动串联起鼓浪屿的如画景致和音乐秘境。用户只需穿戴VR眼镜和手柄，用户便能以“家庭音乐会嘉宾”的身份，跟随“时空领航数字人”穿越时空隧道，用第一视角开启寻找、守护、修复的时空交错之旅。



图12 “琴岛奇遇” XR大空间

案例2：吉米绘本元宇宙戏剧

2024年10月由域上和美出品，携手桑达股份、中国电子云等共同打造全球首个 5G-A VR 大空间沉浸剧场应用。此次 VR 剧场项目以吉米绘本的五大经典之作《我的世界都是你》为灵感来源，以宠物、失去与离别为主题，讲述了一个温暖治愈的故事。通过运用前沿的元宇宙数字技术，在 720° 全方位视角的呈现下，吉米绘本中的那个温暖而治愈的世界被重构得栩栩如生。观众仿佛置身于一个充满奇幻色彩的虚拟空间中，随着剧情的推进，在飞升、巨物等新奇的视觉元素和震撼的音效共同作用下，观众不仅享受到了奇幻绚丽的视觉冲击，更在潜移默化中领悟到了关于生命、爱与成长的深层哲思。



图13 吉米绘本元宇宙戏剧

案例3：《无限TheInfinite》XR 太空沉浸展

《无限TheInfinite》VR太空沉浸展的内容改编自艾美奖获奖作品《太空探索者：国际空间站体验》，由 Felix&Paul和PHI工作室联合创制。VR体验内容时长约60分钟，凝聚了250小时国际空间站VR影片拍摄成果，1:1精心再现国际空间站真实场景。参与者通过VR，即可享受个性化探索路径，“浸”入距离地球400公里的真实太空生活。2024年6月30日，《无限TheInfinite》首次引进中国，于上海新世界大丸百货举办展映。在咪

咕、华录超清、上海拓明等单位支持下，近1000平方米展览区可实现超150人高定位精准度的自由漫游。自2021年蒙特利尔全球首展以来，《无限TheInfinite》已吸引全球50多万观众参与体验，并在持续不断突破新高。



图14 《无限TheInfinite》XR太空沉浸展

案例4：《一梦入大唐》VR大空间

2024年7月3日，NOLO、野子科技联合推出《一梦入大唐》VR大空间。内容制作团队根据历史文献记载和考古挖掘实证，以虚拟现实技术同史料史实创新结合，用户只需佩戴好XR设备便能瞬间穿越至盛唐时代，跟随数字导游的脚步，身临其境地探秘尚未被挖掘的唐太宗昭陵、围观大唐名将震撼的贴面打斗、俯瞰长安城的繁华盛景，开启一场穿越千年的大唐漫游之旅。《一梦入大唐》获CCTV-1、CCTV-13、北京电视台及人民网、北京日报等多家媒体专访报道，吸引超百名达人探店。2个月内体验用户超过1万人次，美团店铺“一梦入大唐”霸榜“北京VR好评榜第一名”，猫眼评分高达9.9分（满分10分）。10月11日，NOLO联合咪咕携《一梦入大唐》亮相第12届中国移动全球合作伙伴大会，吸引大量参会人员进行游玩体验。



图15 《一梦入大唐》VR大空间

案例5：《钢铁突袭》游戏竞技VR大空间

咪咕数智竞技科技中心设有《钢铁突袭》游戏竞技VR大空间项目，基于空间定位、虚拟现实、动作捕捉等

技术，玩家在配备专业的穿戴设备后，立即变身持枪战斗人员，可以组成小队，进行2对2的激烈竞技对抗，充分考验自己的战斗技能和团队合作精神。



图16 “钢铁突袭”游戏竞技VR大空间

2. XR车载场景

智能座舱的发展是全球战略规划，是否具备智能座舱已成为当前用户买车的关键需求。在智能座舱的相关技术当中，车载 XR 又是其关键的能力之一。车载 XR 为驾驶者和乘客带来更丰富直观的信息交互体验，车载 XR 技术还与车机、手机联动，打造多终端融合体验。例如，Rokid 与理想汽车合作发布的 Rokid Max 支持 L 系列车型，实现了观影、游戏、线上办公等功能的投屏体验。星纪魅族发布的 MYVU 智能 AR 眼镜通过 Flyme Link 与魅族手机互联，实现了数据的无缝流转，打造了多终端融合体验，Hololive 公司的技术将扩展现实和车辆运动轨迹参数进行融合，提供全新的未来车载娱乐体验。大众汽车也通过应用扩展现实 XR 技术制作了创新短片，以沉浸式的交互方式展示其纯电动 SUV。小象智能（ELEJO）是国内首家提供汽车与 XR 融合且可实车体验的科技公司，其技术可以消除用户乘车的眩晕感，提供沉浸式虚拟视界。随着光学器件、显示技术、传感器等硬件技术的不断进步，车载 XR 设备将变得更加轻便、舒适，同时提供更高的分辨率和更广阔的视场角，增强用户的沉浸感，让更多开发者投入到车载 XR 内容的开发中，推动内容生态的丰富和多样化，覆盖游戏、教育、娱乐、专业培训等多个领域。5G-A 网络的超低时延、超高带宽将保障 XR 设备性能在移动场景下的最大发挥，预示着未来 5G-A 与 XR 结合无限可能。

案例 1：江西移动车载 XR 示范点

2023 年中国移动、中移虚拟现实、华为、小象智能、中汽中心以及信通院在南昌打造了车载 XR 样板点，共同组成 XRMA 产业联盟并联合发布了《5G-A 车载 XR 用户体验白皮书》，白皮书对车载 XR 用户体验测评标准升级，构建了车载 XR 用户体验 5G-A 网络评估方案。目前白皮书已经应用到车载 XR 行业中。



图17 江西移动车载XR示范点

3.1 企业场景

移动超高清技术也在企业用户市场持续发力，XR产品深入到行业核心痛点和细节需求上，并针对性地探索解决办法，在医疗、教育、文旅、工业制造、赛事演绎、大型活动等实际应用场景中已拥有成熟的通用行业解决方案。

1. 医疗场景

在医疗领域，XR技术用于手术模拟、远程医疗和患者康复训练。通过VR/AR技术，医生可以在虚拟环境中进行手术演练，提高手术成功率。同时，患者可以通过沉浸式体验进行康复训练，加速身体恢复过程。

案例1：VR外科手术

温州医科大学附属第一医院团队利用VR头戴设备，成功完成了两例肠癌手术，将VR技术与微创手术机器人相结合。不同于传统的手术模式，“VR头戴式显示设备+微创手术机器人”的“黄金组合”，将进一步解放主刀医生的头颈部及手部运动，大大提升了手术的自由度和操作的精准度。同时，得益于头显设备中的多功能操作界面，术中主刀医生仅需做几个简单手势，便能随时调出患者的信息及术前CT影像，包括大血管的3D重建图像，结合术中的实时画面，进行精准的血管解剖、淋巴解剖和层面解剖。这种有实时参照的“开卷”手术，让医生们在不离开手术现场的情况下更加安全、高效地作出临床判断和决策。



图18 VR外科手术

2. 教育场景

XR技术在教育和职业培训领域有着广泛应用，XR技术在企业培训中的应用已经逐渐成为一种重要的趋势，它可以模拟产品的机制，帮助员工理解组件如何与机器配合，并提前识别潜在风险。此外，XR技术还能够通过3D模型和虚拟环境，使培训更加直观和高效，从而提高信息处理速度和记忆持久性。

案例1：吉林水利电力职业学院沉浸式互动学习教室

在教学中经常会使用沉浸式互动学习教室，其硬件产品选用了DPVR E4及DPVR P1 PRO，搭配了大朋VR自主研发的星链解决方案，该解决方案能满足近48人的课堂教学。老师使用DPVR E4进行专业内容操作示范，通过星链技术实时同屏到48位学生佩戴的DPVR P1 PRO中，学生能以教师的第一视角沉浸体验实操的全过程，帮助学生了解复杂的建筑专业知识，更直观地学习装配式施工相关操作。



图19 吉林水利电力职业学院沉浸式互动学习教室

3. 文旅场景

XR通过虚拟重建3D场景和交互式数字故事讲述等方式，使用户与文化遗产进行互动，从而增强展示效果。XR技术还可以创建虚拟副本，让游客无需亲临现场即可体验目的地，从而减少对文化遗产的物理压力，也为游客提供沉浸式的旅行体验。

案例1：咪咕“鼓浪屿元宇宙”

厦门市人民政府联合咪咕，依托先进通信、数字孪生和云渲染等技术，采用“咪咕筑境”“咪咕绘境”融合创新引擎，打造国内首个以鼓浪屿为原型的数字空间“鼓浪屿元宇宙”。针对鼓浪屿 1.88 平方公里的自然景观和人文建筑进行毫米级 1:1 数字孪生建模，100%真实还原鼓浪屿全岛的实景面貌。同时，利用动态天气系统、实时光线追踪技术、lumen（动态全局光照）技术等多项前沿科技为“鼓浪屿元宇宙”塑造了一个跨越时空的沉浸式元宇宙世界。此外，“鼓浪屿元宇宙”中还提供了丰富的实时互动玩法，用户通过专属 3D 数字人形象，即可与来自不同地区的陌生人同场竞技、模拟弹奏乐器、互加好友，进一步加深人与人、人与数字世界的交互方式。

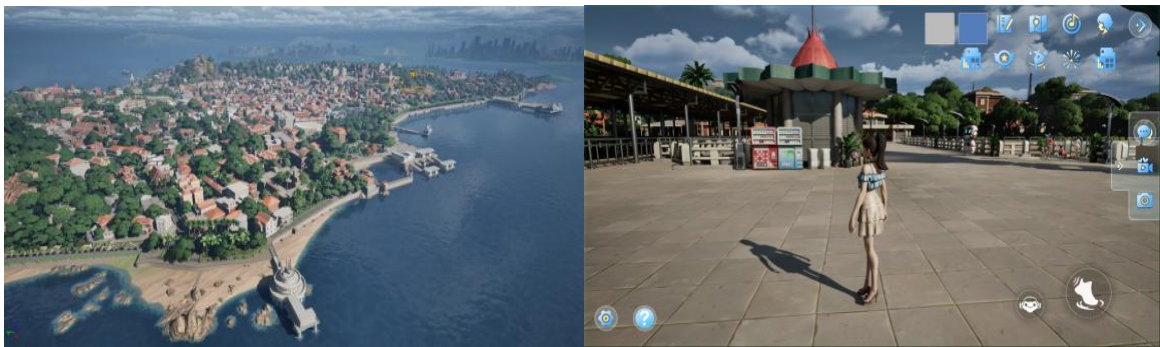


图20 鼓浪屿元宇宙空间

案例2：跟着德爷闯东非

《跟着德爷闯东非》是PICO首部VR互动纪录片，本项目联手全球纪录片领域头部品牌DISCOVERY，与明星探险家德爷（Edward James Stafford）共赴非洲实地拍摄，以实拍+虚拟互动的方式，通过“钻木取火”结合野外生存的定制化的互动设计，让观众以第一视角摄影师的身份陪伴德爷一起冒险，近距离观察野生动物，体验非洲大草原的野外生存之旅。



图21 跟着德爷闯东非

案例3：咪咕“故宫文物南迁纪念展”

2023年12月28日，恰逢故宫文物南迁90周年之际，由故宫、人民日报《国家人文历史》杂志社主办，中国移动咪咕承办的“国宝的长征—故宫文物南迁纪念展”在上海市历史博物馆拉开帷幕。此次展览以故宫文物

南迁史料及研究成果为主要内容，创造性地将室内展厅与室外景观庭院相结合，室内展览分为“国之瑰宝 烽烟南渡”“沪上四载 古物重光”“文脉赓续 生生不息”三个单元，室外展览则以南迁文物箱为造型打造一系列“文物影像屋”。展览通过数智化、沉浸式的呈现技术手段再现故宫文物南迁的历史记忆，带领观众走进一场跨越90年的时空对话。



图22 故宫文物南迁纪念展——文物影像屋

案例4：咪咕“敦煌数字空间AR智能导览”

敦煌研究院、咪咕和Rokid等单位基于AI引擎、图片识别、3D场景呈现等新技术，联合打造“敦煌AR智能导览”作为展览唯一应用的数智化导览系统，通过对莫高窟第285窟进行1:1孪生复刻及数智导游解说，重现千年敦煌的绚烂文化。游客穿戴AR设备便可以360°全景式、沉浸式地“走入”敦煌壁画中，数字导览员“敦敦”指引观众了解莫高窟的前世今生，助力敦煌开启数字化游览新篇章。

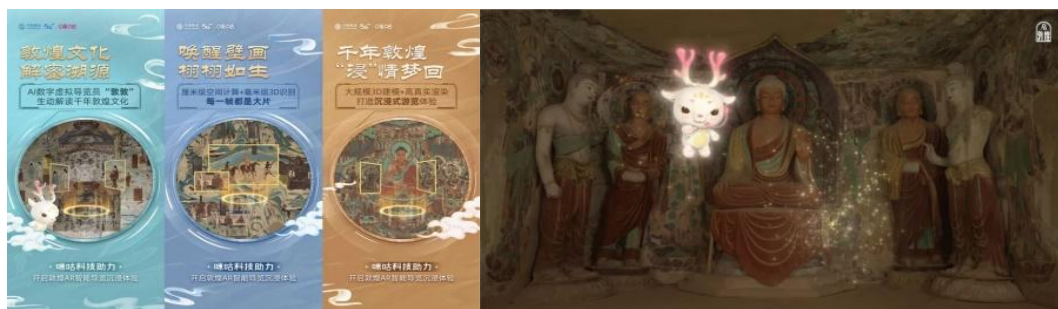


图23 敦煌数字空间AR智能导览

4. 工业制造

XR技术已在工业制造领域得到了广泛应用，包括远程研究、维护和生产制造等环节。通过XR设备，企业可以实现虚拟设计、生产过程模拟、质量管控等，从而提高开发效率，减少成本和时间，同时XR技术为制造业提供了更加直观且实时的设计以及更加仿真的环境，工程师可以利用XR技术创建虚拟原型，快速进行设计的验证和修改，从而缩短产品开发周期、减少开发成本。

案例 1：美的 AR 设备巡检项目

杭州灵伴科技和美的集团的AR巡检项目在设备巡检过程中，作业人员配备AR设备，用于巡检作业中的仪表识别、数据记录、识别状态获取、作业录像留底等功能，提升巡检作业效率。同时为作业人员的眼镜安装远程协作功能，在现场工作人员遇到困难时，专家可以远程协助解决问题。



图24 美的AR设备巡检项目

5、赛事演绎

在赛事演绎方面，XR技术可以为用户提供沉浸式的观赛体验，让他们仿佛置身于比赛现场。通过虚拟现实技术，用户可以看到比赛的全景，感受到现场的气氛和观众的热情，此外，元宇宙还可以提供多角度的观赛视角，让用户可以自由选择自己喜欢的视角观看比赛。

案例1：Pico 2022年世界杯直播

用户支持的球队视角360°3D 7680x7680@30fps直播（双180°3D拼接而成保证正前方和正后方的清晰），感受360°沉浸式球迷们的包裹感，与嘉宾KOL、足球宝贝零距离交流互动，虚拟互动场包厢视角180°3D 7680x3840@30fps直播，从二楼包厢视角观看比赛并可俯视酒吧全景，可观看比赛用时与其他线上用户实时交流互动。



图25 Pico 2022年世界杯直播

案例2：2023年杭州亚运会AR开幕式

作为2023杭州亚运会开幕式5G元宇宙沉浸互动合作伙伴，咪咕突破时空、地域、形式限制，创新打造亚

运会史上首场“数实同屏、瞬时定位、实时互动”的数实融合AR互动节开幕式。通过“放飞亚运华灯”“亚运吉祥物AR互动”“寄送专属亚运明信片”“助燃亚运火炬”四大沉浸式数实融合互动环节，为超7万人的场内观众提供更富趣味性的开幕式体验。



图26 2023年杭州亚运会AR开幕式

案例3：咪咕 “十六届动感地带世界杯音乐盛典”

自2007年首度举办至今，咪咕汇一直是风靡全国、长盛不衰的音乐盛典IP。十六届动感地带世界杯音乐盛典·咪咕汇基于高精度建模、实时渲染、全息投影等前沿技术，以VR/AR/MR+5G+4K舞台的呈现方式，打造了全球首个“全场景数实融合世界杯元宇宙音乐盛典”。全息虚拟舞台效果伴随不同歌曲意境进行改变；周杰伦、刘畊宏、徐梦桃等明星与专属数智人（“橙络络”“刘教练”）同台表演；广大观众通过至臻盛典视角、经纪人视角、神秘保镖视角、随身娱记视角等17路视角享受身临其境的沉浸式全场景演艺观赏体验。



图27 咪咕汇元宇宙舞台

4. 移动场景扩展现实面临的挑战

4.1 沉浸式体验

沉浸式体验的核心在于通过多种感官刺激使参与者完全融入特定的情境或环境中，这种体验不仅限于视觉和听觉，还包括触觉、嗅觉等多种感官的综合运用，沉浸式体验的关键在于“临场感”，即使用者感觉身临其境，所见皆“真”，并且能够与之自然互动。

1、视觉：由于 XR 眼镜屏幕与我们的眼睛距离非常近，因此 XR 的视觉品质需要屏幕像素能够达到与真实世界无限接近的程度，从而做到“以假乱真”的真实感，高分辨率的屏幕同时也去除了“纱窗效应”（screen door effect），即当像素密度不足时画面在人眼中产生虚影的现象，就像通过纱窗看外面的世界。人类通过双眼以 3D 形态观察世界，这就需要立体显示器为场景中的物体显示适当的景深，从而实现 3D 效果。人眼拥有接近 180° 的视场角（field of view，简称：FOV），而能清晰成像的面积也有将近 120°，因此 VR 设备的 FOV 越接近这一数值，越能提供沉浸式的体验。视觉上的考验还远不止如此，更加困难的是从人体运动到显示屏刷新之间的延时（motion to photon，简称：MTP）问题。通过实验，MTP 延时超过 20ms 将使沉浸式体验显著下降，同时用户也会产生眩晕感。此外，显示屏的刷新频率越高，所要求的 MTP 延时越低，例如一块 60Hz 的屏幕的合理延时需少于 17ms，而一块 90Hz 的屏幕则需低于 11ms。但当前很多 XR 设备在分辨率、延迟、舒适度等方面或多或少还存在一些问题，需要技术不断改进和完善，以提高用户的体验和满意度。

2、听觉：为了使视觉与听觉得到完全同步，我们需要高保真音频和 3D 环绕声，高保真的音频可以在采样频率和采样位数上与人类听力相匹配。我们处于三维空间当中，声音是 3D 形态的，同样在 XR 场景中，3D 音频也必不可少，逼真的 3D 定位音频可以在头部转动时动态调整，这就意味着我们可以在虚拟现实环境中，通过声音来辨别声源方向和距离，做到与真实生活别无二致。尽管空间音频技术正在发展，但目前商用的 HRTFs（头相关传输函数）数据库的精度尚未达到与人耳听辨能力相媲美的水平。此外，每个人的人体工学参数和心理声学系统都不同，甚至会随着年龄增加而变化，这使得实现高度逼真的听觉体验更具挑战性。

3、交互：直观交互实际上是人类的第二本能，没有什么比转动头部环顾四周更自然的方式了，依赖于认知技术，我们还可以使用手势或语音，在虚拟空间进行控制。六自由度（6-DOF）精准运动跟踪不仅可以实现上下左右的基本模拟，还可以在不断变化的虚拟现实环境中，实现前进和后退。受益于近年来的技术发展，虚拟现实已经摆脱了早期笨重的有线头戴式显示器和控制器。移动平台开始采用系统性解决方案和基于 SoC 的定制设计专业引擎，提供经过终端到终端优化后的高效异构运算解决方案，将 XR 变成现实，手部跟踪、眼动追踪和语音交互等技术是当前 XR 体验中的关键挑战。手部跟踪需要精确识别用户的动作，而眼动追踪则可以更符合人体工程学，提高界面的可用性。尽管多模态交互技术已经获得了很大的进步，但其复杂性也带来了不少挑

战。例如，手势识别的复杂性、面部表情和注视的多模态融合难度较大，这需要进一步的技术突破。此外，眼手交互虽然提高了交互效率，但也存在学习成本高、鲁棒性不足等问题。

4.2 丰富的内容

丰富的内容是吸引客户使用 XR 的关键因素，根据多方面的证据，XR 行业的发展和用户体验的提升都依赖于高质量的内容。内容是 XR 行业的重要组成部分，也是吸引用户的关键因素，高质量的游戏、电影和教育等 XR 内容可以为用户提供丰富的娱乐和学习体验。此外，好的 XR 体验必须具有沉浸感和吸引力，不仅要能够吸引消费者的注意力，还要能提供价值并增强消费者与品牌的互动体验。同时，XR 设备本身无法脱离内容来提供功能价值，消费者购买 XR 设备主要是为了体验设备上的内容或应用，因此完善的内容是促进 XR 设备销量的重要因素。但实际上 XR 内容数量有限，例如截至 2022 年 6 月，Oculus Rift 平台上的 XR 内容数量仅为 1392 款，而 Quest 系列硬件累计出货量约 2000 万台，但内容数量仅 400 余款，这种内容数量上的不足限制了消费者的选择和体验。导致内容缺失的有以下几个关键问题：

- 1、高质量内容稀缺：尽管有大量内容存在，但高质量的 VR 内容仍然稀缺。上榜作品榜单上仍以老牌 VR 游戏为主，近年来现象级重磅游戏较为匮乏，这表明市场上缺乏能够吸引大量用户并产生持续兴趣的高质量内容。
- 2、开发成本高且收益低：部分内容复杂、画面精致的虚拟现实作品收益不能覆盖制作成本，削弱了开发团队的创作热情，制约了内容创作进度，降低了 VR 游戏效果，高额的硬件购置成本也影响了终端用户的需求。
- 3、人才紧缺和传统思维：部分虚拟现实内容开发企业由电脑端游戏团队、移动端游戏团队转型而成，其作品设计思路受制于传统内容开发模式，导致作品画面复杂、操作难度大、用户学习成本过高等问题。
- 4、市场环境和用户黏性不足：XR 用户黏性不足、场景较少是限制 XR 销量持续成长的重要因素，行业亟待新玩家破局，缺乏优质内容生态供给成为老生常谈之言。

4.3 便携和续航

移动场景下使用 XR 最关键的诉求是方便携带和续航能力。XR 设备通常需要集成高性能的 CPU、存储和传输组件，这导致了设备的重量较重，用户难以长时间佩戴，目前市面上主流的 XR 头显重量大多在 500g 左右，甚至部分设备达到了 650g，这种重量不仅影响了用户的佩戴舒适度，还可能导致长时间使用时出现颈部疼痛。此外，XR 设备的重量与其性能之间存在一定的不平衡。随着 XR 设备朝着高自由度和高沉浸度的方向发展，设备中的追踪方案和算力系统被集成到头显中，使其更加笨重。尽管有研究和尝试通过轻量化设计来改善

这一问题，但整体上设备仍然面临重量过重的问题。另外，XR 设备的重量还与佩戴者的体验密切相关。较重的设备会增加头部负担，长时间佩戴后容易造成疲劳和不适感。因此，减轻 XR 设备的重量是提升用户体验的重要方向之一。XR 设备的重量问题不仅影响用户的佩戴舒适度，还限制了其使用时长和体验质量，未来的发展趋势是朝着更轻便、更舒适的方向努力，以提高用户的接受度和使用时长。

续航是另外一个用户最关心的问题，想要 XR 设备可以长续航，目前还存在一定的困难。XR 设备需要高亮度和高性能的显示效果，这导致了较高的功耗。例如，Pico 4 XR 头显由于采用了 Pancake 光学方案，需要很高的亮度，同时 XR2 芯片性能不足，为了实现较好的渲染效果，必须使用 CPU 高频运行，这进一步增加了功耗。高性能的计算需求也使得 XR 设备的电池续航时间受到限制。XR 设备的重量和散热问题也同样影响了续航能力。一体式 XR 设备由于集成了强大的计算能力，导致其体积较大且重量较重，这不仅影响了佩戴体验，还可能导致设备过热，从而缩短电池寿命。据用户反馈显示，XR 设备在高帧率内容运行下，电量只能维持约 1~2 个小时，这显然无法满足长时间使用的需求。

5. 技术标准化与产业政策建议

5.1 关键技术标准化及发展方向

XR技术的标准化工作正在全球范围内积极推进，以应对日益增长的技术需求和行业挑战。

1、动作捕捉：世界超高清视频产业联盟制定了 T/UWA 022.1-2023、T/UWA 022.2—2024 和 T/UWA 022.3-2023 三项标准，分别涉及虚拟现实应用中的人体动作捕捉系统的通用技术要求、惯性动作捕捉系统软件接口和光学动作捕捉系统软件接口。这些标准规定了虚拟现实人体动作捕捉系统的功能和性能标准，以及惯性系统的软件接口要求，旨在指导相关系统的研发、应用和维护。在国际层面，ISO/IEC JTC 1 发布了类人动画系列标准，包括 ISO/IEC 19774-1 和 19774-2，分别定义了人形数据结构和动画运动数据，用于动作捕捉数据的投影和动画生成。中国提出的《信息技术 面向虚拟现实应用的人体动作捕捉系统》国际标准也计划纳入这一体系。实景拍摄和动作捕捉作为扩展现实内容生产的关键，涉及的内容采集流程、数据格式和音视频设备等都是未来标准化工作的重点，这些标准的制定将提升内容生产效率，促进行业的规范化发展。

2、三维重建：TC 28/SC 24（全国信标委计算机图形图像分委会）组织起草《信息技术 大规模场景多视图三维重建系统技术规范》（计划号：20240831-T-469），给出了基于多视图影像的大规模场景三维重建系统的总体原则、功能要求、性能要求和测试方法等。

3、XR 制作系统：世界超高清视频产业联盟发布了《XR 虚实融合制作系统 技术要求与测试方法》标准（T/UWA 016-2022），该标准涵盖了 XR 虚实融合制作系统及其子系统（LED 大屏系统、拍摄系统、目标跟踪系统、渲染系统）的技术要求与测试方法，该标准为指导 XR 虚实融合制作的规范开展提供了明确的技术指导，也为开展面向后续虚拟影棚、虚拟演播室等具体应用场景的标准化工作奠定了基础。

4、编解码：国际上和国内在全景视频、自由视点视频、3D 点云、3D Mesh 和三维声（Audio Vivid）等超高清音视频技术领域的发展和标准化工作正快速推进。国际组织如 ITU-R SG6、MPEG、3GPP、JPEG 和 IEEE 等都在制定相关标准，以规范这些技术的发展。国内方面，中国的 AVS 工作组也在推动 VR 视频编码等国家标准的制定。这些标准旨在提升视频质量、优化传输效率，并支持 5G 网络、有线电视网络、IPTV 和互联网传输分发等新的应用场景。未来，高效的球面-平面的投影、高效的片划分、提高并行处理速度、实现用户姿态变化快速响应及视角自适应动态切换等将是未来全景视频编码技术的重要发展方向。更高效的补丁块生成、深度图生成方法、兼容更多的视频编码器、基于率失真优化的联合比特分配、面向全景内容的非同构视图选择、支持更宽的深度动态范围并实现占用误差校正、降低视点切换的延迟等将是未来自由视点视频编码的技术的发展方向。更高效的几何表征方法及属性预测技术、更灵活的编码速度配置，支持多种离线转码、云游戏、视频直播和视频会议等延迟场景，兼容各终端设备，实现场景自适应，提高编码速度（帧间并行加速、预分析和后处理加速），智能码率控制等是点云编码技术发展方向。

5、传输：移动场景需支持多种视频数据传输，强调实时性和互动性。3GPP SA4 正研究 5G 标准项目，如 5G_RTP、iRTCW 等，以优化沉浸实时业务的媒体。IETF 和 W3C 已采纳 WebRTC 为标准，并研发下一代标准，探索边缘计算影响及用例，整合网络质量监测。为支持移动端超高清媒体传输，研究扩展和优化媒体传输协议（如 RTP、WebRTC）、功能组件，并提供灵活的沉浸媒体数据格式传输/访问机制，以提升效率、减少开销、增强沉浸体验，满足多元业务场景。

6、业务平台：音视频业务平台依托基础算力网络，旨在提供全面的视频基础业务能力，并向上层应用服务与终端开放强大的应用能力，涵盖视频平台、云渲染平台及 AI 算法平台等多元化能力以及高效的管理运营平台。

鉴于移动场景的丰富性与开放性特点，建议深入探究业务平台的架构设计、功能规范、接口标准、技术实施流程及网络配置需求等核心技术指标。

7、终端显示：移动场景下，终端显示设备主要以头戴式为主，全国信息技术标准化技术委员会计算机图形图像处理及环境数据表示分委会（SAC TC 28/SC 24）归口国家标准 GB/T 38259-2019《信息技术 虚拟现实头戴式显示设备通用规范》规定了三种形式的虚拟现实终端设备的基本要求、试验方法、检验规则及标志、包装、运输和贮存等。全国音频、视频及多媒体系统与设备标准化技术委员会(SAC/TC242)在研行业标准《超高清虚拟现实显示设备通用规范》（计划号 2019-1104T-SJ）规定了超高清虚拟现实显示设备的性能要求和测试方法。眼镜式的设备具有其独特的技术特性和舒适度要求，这些设备通过将虚拟信息与现实世界巧妙结合，为用户提供增强的感知体验。随着技术和产业的发展，增强现实设备正逐渐成为未来标准化的重要方向之一。

8、评测标准：XR 移动场景具有沉浸式体验、高交互性、便携性与灵活性、实时渲染与高效传输以及跨平台兼容性等特点，对视频质量、交互性能、兼容性等提出更高要求。未来评测标准化主要针对分辨率、帧率、色彩还原度、对比度与亮度的评测，确保画面清晰、流畅且真实这几方面展开。同时对实时性、互动性和稳定性评估，保障用户能实时、便捷地与视频内容互动。测试视频在不同移动设备和操作系统上的表现，确保跨平台流畅播放；评估网络带宽需求、延迟与抖动，确保低延迟、稳定的观看体验以及监测视频播放时的能耗和发热情况，保障设备续航和舒适度。

5.2 产业政策建议

移动场景扩展现实超高清技术的发展不仅仅是技术本身的标准化及创新，更离不开产业政策的支撑，截至目前国家已经出台一些支撑政策，如 22 年工信部等 5 部委发布的《虚拟现实与行业应用融合发展行动计划（2022-2026 年）》、23 年工信部 7 部委发布的《关于加快推进视听电子产业高质量发展的指导意见》，但为了加速移动场景扩展现实超高清技术的高质量发展，建议出台更多方向的产业政策。

1、加强技术研发与创新：支持超高清视频核心技术研发，如高分辨率成像、高带宽实时传输、HDR 显示兼容等，并推动这些技术在移动场景下的应用，推动 XR 终端的混合现实技术发展，鼓励企业在交互形式和信息呈现上进行创新，特别是在工业制造领域。

- 2、产业链协同发展：促进超高清视频产业链上下游的协同，形成完整的产业生态，加强 OTT 行业与内容提供商、服务运营商和设备制造商之间的合作，共同提升服务能力和用户体验。
- 3、市场与应用推广：通过政策支持，推动超高清视频内容的制作和分发。支持 4K/8K 超高清电视频道建设，并在移动场景下推广应用，鼓励 MR 技术在工业、医疗、教育等领域的应用，通过实际场景的落地来推动技术成熟和市场接受度。
- 4、质量认证与标准制定：建立和完善移动场景超高清视频产业的标准体系，包括采集、制作、传输、呈现、应用等全产业链的标准，推动 MR 终端和 OTT 设备的质量认证，确保产品和服务的质量，提升消费者信任。
- 5、人才培养与引进：加强超高清视频产业人才队伍建设，培养高端、复合型领军人才，并鼓励产学研合作。支持高等学校相关学科设立，引导职业学校培养产业发展急需的技能型人才。
- 6、国际合作与交流：推动国际合作，引入国际企业，建立长效合作沟通机制，加强项目合作、产学研合作，构建互利共赢的产业生态。支持国内企业参与国际标准制定，提升国际影响力和话语权。
- 7、政策激励与约束：通过设立产业投资基金等方式，支持超高清视频产业创新发展，并引导多方资本参与，强化牌照方在 OTT 行业的管理定位，推动行业的有序、高质量发展。
- 8、公共服务平台建设：建设超高清视频产业支撑服务平台，提供标准制定、评测认证、视频制作、版权交易等支撑服务，支持建设超高清视频内容制作生产基地和内容集成平台，丰富超高清视频内容资源。

6. 发展趋势展望

移动超高清市场正处于快速增长期，预计未来几年将成为主流选择，XR 技术的融合为超高清带来了更广阔的发展空间，为用户带来更加丰富、真实的沉浸式体验。

- 1、移动通信技术的推动：随着 5G 和 5G-A 技术的发展，其高带宽低时延特性显著改善了 XR 高清视频的传输和体验，高带宽特性能够解决高清 XR 内容的无线传输问题，使用户可以流畅地观看超高清 XR 视频，低时延特性有效减轻了因画面延迟产生的眩晕问题，这对于提升用户的沉浸式体验至关重要。5G 和 5G-A 技术支持边缘计算，可以将计算资源卸载到边缘网络，从而减少延迟并提高 XR 应用的性能，这种端边协同分离渲染技术使得 XR 用户能够借助 5G 高速率低时延的网络传输来协同边缘云侧进行实时渲染，同时结合 XR 终端侧的本地优化渲染，提供低时延沉浸式无界 XR 体验。5G 和 5G-A 还支持多模态传输和系统容量提升等高级功能，进一步优化 XR 体验。例如，通过支持额外的缓冲区大小表和非整数 DRX 周期性，5G 能够更好地匹配帧率需求，

并减少量化误差，这些技术创新不仅提升了 XR 设备的感知交互能力，还推动了 XR 产品形态和功能的进一步优化。

2、内容生态的丰富：超高清视频作为 XR 设备的刚需型应用，正在成为推动用户体验提升的重要因素。通过 MV-HEVC 编码提高存储效率的技术和空间视频形式受到用户的好评，有望成为杀手级应用。此外 AIGC 的快速发展，让 UGC（用户生成内容）百倍、千倍的增长，为 XR 产业带来了新的增长动力，丰富了视频内容并提高了用户的参与度。

3、硬件生态的发展：智能手机与 XR 设备的互联正在推动 XR 设备的销量增长，XR 设备可以通过有线或无线的方式与智能手机互联，有线连接通常通过电缆将 XR 头显连接到路由器或交换机，或者先连接到一个 XR 专用本地设备，再由该本地设备连接到路由器和交换机，最终连接到 XR 服务器。无线连接则依赖于 5G 等高速网络技术，使 XR 设备能够实现更灵活的使用。随着 XR 技术的普及，如视频桥接芯片等相关硬件的需求也在增加。智能手机与 XR 设备互联的技术细节和实现方式涉及多个方面，包括连接方式、硬件优化以及软件支持等。高通公司推出的第二代骁龙 XR2+ 平台，旨在加速 MR 体验的发展，这表明智能手机在处理 XR 内容时需要具备高性能低功耗计算能力。终端侧 AI 技术的应用也使得智能手机能够在本地执行复杂的指令，从而减少对外部计算资源的依赖。MineXR 项目（设计挖掘工作流程和数据分析平台，旨在帮助研究人员收集有关个性化扩展现实布局的数据，以了解未来 XR 的使用情况并开发自适应）的研究结果表明，设计模块化 UI 元素、适应用户空间和活动的 XR 界面、提供外围功能的连续访问等设计指南有助于提升 XR 体验的个性化和有效性。智能手机作为 MR 眼镜的配件使用，可以借鉴 VR 盒子的成功经验，通过智能手机的强大计算能力和显示能力来增强 XR 设备的体验。智能手机可以录制空间视频并在 XR 设备上播放，不同品牌的 XR 设备之间也可以互相播放录制的内容，这促进了硬件生态的互联。

4、技术创新：3GPP TR 26.928 版本 18.0.0 发布了关于 XR 的最新标准，涵盖了 XR 场景生成、渲染、媒体解码等方面的内容，并提出了基于网络的媒体处理接口，以支持各种类型的媒体处理。

1) QoS（服务质量）和 QoE（用户体验）参数：该版本详细介绍了与 XR 相关的 QoS 和 QoE 参数。

2) 潜在标准化需求：报告分析了现有 5QIs（服务质量指标），并提出了潜在的标准化领域。

3) 网络渲染：包括基于栅格的分割渲染、通用 XR 分割渲染以及 XR 分布式计算等内容。

4) 流量特征和模型：总结了 XR 服务的流量特征和模型。

5) XR 用例：如 3D 图像消息传递、AR 共享、沉浸式 6-DOF 流媒体和情感流媒体等。

- 6) 三维向量和 XR 空间概念：定义了三维向量，并介绍了 XR 参考空间的概念，用于与用户物理环境建立空间关系。
- 7) 媒体格式和传输技术：涵盖了音频/视频格式的编码表示、几何数据、场景组成和描述、存储和云访问格式、支持任何媒体的传输协议、解码、渲染和传感器 APIs、生物识别和情绪元数据定义与传输、故事线的无缝拼接和平滑过渡以及室内空间信息的存储和共享格式。
- 8) XR 会议应用程序：评估了多个用例，包括 XR 会议应用与 5G 系统架构的映射、支持网络中的媒体处理、6-DOF 元数据框架、3D 对象格式和传输、数据采集和存储、内容交付协议等。
- 9) 增强现实的新形式：讨论了玻璃型 AR/MR UEs，并将其归类为 XR5G-A5（具有独立连接能力，可以直接连接到 3GPP 网络的设备）；同时考虑了主要的技术约束是设备的形态因子和发射功率，如果 XR5G-A5 设备以眼镜的形态因子实现（小于护目镜），并且重量小于 100 克，其蜂窝调制解调器和天线位于面部附近，这将导致 XR5G-A5 设备在发射功率上受到更多限制，可能需要退化为 XR5G-A1（有线的简单 XR 可穿戴设备，电源、XR 处理和网络连接都是由外部源提供的）或 XR5G-A2（无线的简单 XR 可穿戴设备，电源、XR 处理和网络连接都是由外部源提供的）模式。
- 10) 扩展的分割和云渲染处理：基于短期开发的工作，定义了扩展的分割和云渲染处理，并考虑了简单扩展以支持 6-DOF 流媒体。

总体来看，移动场景下的超高清视频发展趋势呈现出技术进步、内容生态丰富、硬件生态互联以及应用场景拓展等多方面的积极变化，这些因素将共同推动了移动场景扩展现实超高清技术的快速发展和广泛应用。

7. 附录

7.1 缩略语

下列缩略语适用于本文件。

AIGC 人工智能生成内容 (Artificial Intelligence Generated Content)

AR 增强现实 (Augmented Reality) DRX 间断性接收 (Discontinuous Reception)

MR 混合现实 (Mixed Reality)

MV-HEVC 多视图-高效视频编码 (Multiview High Efficiency Video Coding)

mMTC 海量物联网通信 (Massive Machine Type Communication)

UI 用户界面 (User Interface)

URLLC 超高可靠性与超低时延业务 (Ultra-Reliable & Low Latency Communication)

VR 虚拟现实 (Virtual Reality)

XR 扩展现实 (Extended Reality)

3GPP 第三代合作伙伴计划 (3rd Generation Partnership Project)

3D 三维 (three dimensional)

5G 第五代移动通信技术 (5th Generation Mobile Communication Technology)

5G-A 第5.5移动通信技术 (5G-Advance)

6-DOF 六自由度 (Six Degree Of Freedom)

参考文献

- [1] GB/T 38259-2019 《信息技术—虚拟现实头戴式显示设备通用规范》
- [2] T/UWA 016-2022 《XR虚实融合制作系统技术要求与测试方法》
- [3] T/UWA 022.1-2023 信息技术 面向虚拟现实应用的人体动作捕捉系统 第1部分：通用技术要求
- [4] T/UWA 022.2-2024 信息技术 面向虚拟现实应用的人体动作捕捉系统 第2部分：光学动作捕捉系统软件接口参数
- [5] T/UWA 022.3-2023 信息技术 面向虚拟现实应用的人体动作捕捉系统 第3部分：惯性动作捕捉系统软件接口标准
- [6] ISO/IEC 19774-1 《计算机制图、图像处理和环境数据表示.第1部分》
- [7] ISO/IEC 19774-2 《计算机制图、图像处理和环境数据表示.第2部分》
- [8] XRMA产业联盟《5G-A车载XR用户体验白皮书》
- [9] Gerhard et al. "Overview of the Multiview and 3D Extensions of High Efficiency Video Coding" (2015, September 11).



UHD World Association

世界超高清视频产业联盟

联系我们：
UWA联盟邮箱： support@theuwa.com
UWA联盟官网： www.theuwa.com